

## Durham Research Online

---

### Deposited in DRO:

11 August 2014

### Version of attached file:

Published Version

### Peer-review status of attached file:

Peer-reviewed

### Citation for published item:

Lyratzopoulos, G. and Abel, G.A. and McPhail, S. and Neal, R.D. and Rubin, G. (2013) 'Measures of promptness of cancer diagnosis in primary care : secondary analysis of national audit data on patients with 18 common and rarer cancers.', *British journal of cancer.*, 108 (3). pp. 686-690.

### Further information on publisher's website:

<http://dx.doi.org/10.1038/bjc.2013.1>

### Publisher's copyright statement:

This work is licensed under the Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>

### Additional information:

---

## Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

**Keywords:** cancer; consultations; delay; diagnosis; general practitioner; referral

# Measures of promptness of cancer diagnosis in primary care: secondary analysis of national audit data on patients with 18 common and rarer cancers

G Lyratzopoulos<sup>1,\*</sup>, G A Abel<sup>1</sup>, S McPhail<sup>2</sup>, R D Neal<sup>3</sup> and G P Rubin<sup>4</sup>

<sup>1</sup>Cambridge Centre for Health Services Research, Department of Public Health and Primary Care, University of Cambridge, Forvie Site, Robinson Way, Cambridge CB2 0SR, UK; <sup>2</sup>National Cancer Intelligence Network (NCIN), 18th Floor, Portland House, Bressenden Place, London SW1E 5RS, UK; <sup>3</sup>North Wales Centre for Primary Care Research, College of Health and Behavioural Sciences, Bangor University, Gwenfro Units 4-8, Wrexham Technology Park, Wrexham LL13 7YP, UK and <sup>4</sup>Wolfson Research Institute, School of Medicine, Pharmacy and Health, University of Durham, Queen's Campus, University Boulevard, Stockton-on-Tees TS17 6BH, UK

**Background:** Evidence is needed about the promptness of cancer diagnosis and associations between its measures.

**Methods:** We analysed data from the National Audit of Cancer Diagnosis in Primary Care 2009–10 exploring the association between the interval from first symptomatic presentation to specialist referral (the primary care interval, or 'interval' hereafter) and the number of pre-referral consultations.

**Results:** Among 13035 patients with any of 18 different cancers, most (82%) were referred after 1 (58%) or 2 (25%) consultations (median intervals 0 and 15 days, respectively) while 9%, 4% and 5% patients required 3, 4 or 5+ consultations (median intervals 34, 47 and 97 days, respectively) (Spearman's  $r=0.70$ ). The association was at least moderate for any cancer (Spearman's  $r$  range: 0.55 (prostate) – 0.77 (brain)). Patients with cancers with a higher proportion of three or more pre-referral consultations typically also had longer median intervals (e.g., multiple myeloma) and vice versa (e.g., breast cancer).

**Conclusion:** The number of pre-referral consultations has construct validity as a measure of the primary care interval. Developing interventions to reduce the number of pre-referral consultations can help improve the timeliness of cancer diagnosis, and constitutes a priority for early diagnosis initiatives and research.

In spite of its potential importance (Richards *et al*, 1999; Thomson and Forman, 2009; Torring *et al*, 2011) there is currently no uniform approach to measuring the timeliness of cancer diagnosis. The time between the first symptomatic presentation of a cancer patient to a general practitioner and their first specialist referral for further investigation (i.e., the 'primary care interval') is a relatively established measure (Allgar and Neal, 2005; Neal and Allgar, 2005; Hansen *et al*, 2011; Weller *et al*, 2012). The number of

consultations with a general practitioner before hospital referral is increasingly being used as a measure of the promptness of cancer diagnosis (Pancreatic Cancer UK, 2011; Rarer Cancer Foundation, 2011; Rubin *et al*, 2011; The Roy Castle Lung Cancer Foundation, 2011; Anonymous, 2012; Lyratzopoulos *et al*, 2012; Teenage Cancer Trust, 2012).

Both measures (i.e., both the primary care interval and the number of pre-referral consultations) have specific strengths and

\*Correspondence: Dr G Lyratzopoulos; E-mail: gl290@medschl.cam.ac.uk

Received 11 October 2012; revised 5 December 2012; accepted 16 December 2012; published online 7 February 2013

© 2013 Cancer Research UK. All rights reserved 0007–0920/13



weaknesses, some of which are shared (see Online Supplementary Information, Table). Exploring whether and how the two measures are associated is important: If strongly correlated, both or either can be used for research and quality surveillance; if weakly correlated, using the number of pre-referral consultations as a measure of timeliness should be discouraged. Importantly, if the number of pre-referral consultations is a valid surrogate of the primary care interval, earlier diagnosis efforts could focus on improving the sensitivity of symptom appraisal by general practitioners during a medical consultation, for example, using clinical decision support tools (Hamilton, 2009; Hippisley-Cox and Coupland, 2012). We therefore examined the association between the number of pre-referral consultations and the primary care interval.

## MATERIALS AND METHODS

We analysed data from the (English) National Audit of Cancer Diagnosis in Primary Care (2009–2010), undertaken as part of the National Awareness and Early Diagnosis Initiative (Cancer Research UK, 2009; Rubin *et al.*, 2011). Data were collected by general practitioners or other professionals in an estimated total of 1170 general practices (~14% of all practices in England) who participated voluntarily. Comparisons with cancer registration statistics indicate good representativeness (Rubin *et al.*, 2011). We analysed data on patients aged 15 years or older with completely observed information on number of pre-referral consultations and primary care interval values from 0 to 730 days. Because of sample size considerations, analysis was restricted to patients with the 18 most numerous cancers.

We calculated the median primary care interval (in days) and inter-quartile range for different numbers of pre-referral consultations. This was done both for all patients and separately by cancer. Box plots were used to further illustrate these data. Given the right-skewed nature of primary care interval data, we calculated the Spearman's rank correlation co-efficient of the two measures. Subsequently, we calculated the receiver-operator characteristic (ROC) area under the curve (AUC), treating the primary care interval as a continuous variable and the number of pre-referral consultations as a binary variable ('three or more' vs 'one or two' consultations—as used by UK policy makers (NHS, 2011)).

Using linear regression we explored whether the association between the two measures varied between different patients groups. Owing to data non-normality, significance testing was based on bootstrapping (1000 samples). See also Online Supplementary Information.

## RESULTS

Among 13 035 patients included in the analysis the median (inter-quartile range) primary care interval was 5 (0–23) days. Most (82%) were referred after 1 (57.5%) or 2 (24.6%) consultations, with median primary care interval (inter-quartile range) of 0 (0–3) and 15 (7–31) days, respectively. Of the 18% that required 3 (9.2%), 4 (3.9%) or '5 or more' (4.8%) consultations, the median primary care interval (inter-quartile range) was 34 (16–64), 47 (27–90) and 97 (46–173) days, respectively (Table 1, Figure 1). Similar patterns were observed for patients with any of the 18 cancers (Table 1). Patients with multiple myeloma and lung cancer had high proportions of three or more pre-referral consultations (46% and 33%, respectively) and long median primary care intervals (21 and 14 days, respectively). Patients with breast cancer and melanoma had the lowest proportion of three or more pre-referral consultations (3% and 5%, respectively) and the shortest median primary care intervals (0 days for both).

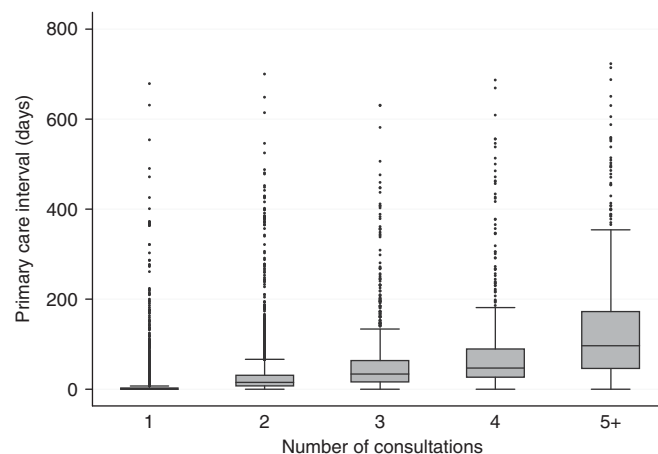


Figure 1. Box plot for primary care interval by category of number of pre-referral consultations (1, 2, 3, 4 and '5 +') for patients with any of 18 cancers ( $n = 13\,035$ ).

For patients with any of the 18 cancers the Spearman's  $r$  was 0.70 ( $P < 0.0001$ ) and the ROC AUC was 0.88 (95% confidence interval: 0.87–0.89) (Table 2). Primary care interval cutoff values of either 17 or 18 days have sensitivity and specificity of 0.80 or higher for predicting pre-referral consultation category.

Using linear regression we found no evidence for interactions between the number of pre-referral consultations and age, sex or ethnicity ( $P = 0.998$ ,  $P = 0.261$  and  $P = 0.691$ , respectively), but strong evidence for interaction with cancer type ( $P < 0.001$ ). Nevertheless, the association was at least moderately strong for each cancer (Spearman's  $r$  range: from 0.55 to 0.77 for prostate and brain cancer, respectively, Table 2). There was evidence for an interaction between the number of pre-referral consultations and different referral types ( $P = 0.016$ ), however the association was strong for any referral type (Spearman's  $r$  range: from 0.67 to 0.72 for routine and '2-week' referrals, respectively, Table 2). See also Online Supplementary Information.

## DISCUSSION

Using data from the English National Audit of Cancer Diagnosis in Primary Care we found that the number of pre-referral consultations has construct validity as a measure of the primary care interval. The association between the two measures was strong for patients with any type of hospital referral and either moderate or strong for patients with any of the studied cancers. Patients with cancers requiring greater number of consultations tended to have longer primary care intervals (e.g., multiple myeloma) and vice versa (e.g., breast cancer).

The findings allow us to gain some insight into the potential reduction in primary care interval that could be achievable by reducing the number of pre-referral consultations. For example, the median primary care interval for patients with 'five or more' pre-referral consultations was 49.5, 62.5 and 81.5 days longer than those with four, three and two consultations, respectively. Such gains in the timeliness of diagnosis could potentially be achievable if patients with different number of consultations present with similar symptoms. These illustrations should therefore be considered to represent the theoretical maximum potential for improvement (in timeliness that could be achievable by reducing the number of pre-referral consultations) given the state of medical knowledge and system factors at the time of the audit period (such as the extent of access to specialist investigation available to general practitioners).

Table 1. Median primary care interval (number of days) and inter-quartile range by pre-referral consultations category (n = 13 035)

Median primary care interval by number of pre-referral consultations category (1-5 +)														
	N	% Of patients with 3 + consultations	Median primary care interval, all patients (n = 13 035/100%)	Inter-quartile range	1 (n = 7499/57.5%)	Inter-quartile range	2 (n = 3202/24.6%)	Inter-quartile range	3 (n = 1195/9.2%)	Inter-quartile range	4 (n = 511/3.9%)	Inter-quartile range	5 + (n = 628/4.8%)	Inter-quartile range
All patients	13 035	17.9	5	0-23	0	0-3	15	7-31	34	16-64	47	27-90	96.5	46-172.5
By cancer														
Multiple myeloma	176	46.0	21	5-55	1	0-12	14	5-33	38	21-64	45	22-82	82	34-129
Unknown primary	136	41.2	14	0-37.5	0	0-3	19	9-30	24	13-42	26.5	16-62	54	37-124
Stomach	237	32.1	14	0-57	0	0-4.5	21	9-42	28.5	16-69.5	66	32-123	127	82-194
Lung	1421	32.8	14	3-40	3	0-10	13	6-27	30	14-56.5	39	23-70	73	42-140
Renal	272	22.1	12	1-37	1	0-12.5	14.5	7-31	30	15-55	48	41.5-85	113	63-185
Prostate	2201	15.2	12	3-28	4	0-13	15	7-28	38	20-77	66.5	36.5-188.5	113	38-181
Lymphoma	590	25.8	9	0-32	1	0-5	14	7-28	32	14-69	42.5	25.5-84	76.5	41-189
Pancreatic	303	31.4	8	1-35	0	0-4	8	4-14	21	13-41	45.5	16-86	93	49-144
Ovarian	324	27.8	8	0-27	0	0-5	11	5-21	39.5	14-54	36	13-49	49	25-71
Oesophageal	500	22.6	7	0-33	0	0-1	20	8-37	33	21-56	76	43-124	119.5	61-245
Colorectal	1999	21.3	7	0-32	0	0-3.5	18	7-40	39	15-75.5	56	30-114	122	56-203
Leukaemia	358	17.6	6.5	0-23	1	0-6	15	7-37	33	13-52	32.5	9-66	25	14-98
Oropharyngeal	178	21.9	6	0-27	0	0-3	17.5	12-33.5	28	18-57	42.5	29-57	64	35-116
Brain	159	21.4	4	0-19	0	0-1.5	7	5-16	25.5	5-64	30	22-62	74.5	49-166
Bladder	721	14.6%	4	0-18	0	0-3	12	7-22.5	22	16-61.5	59	43-141	153.5	73-297
Endometrial	358	9.8	1	0-19	0	0-2	21	8-44.5	41	19-63	70.5	26.5-96.5	258	164-349
Melanoma	735	5.4	0	0-6	0	0-1	30	11-53	37.5	22-95	42	31.5-106.5	100	67-141.5
Breast	2367	2.9	0	0-1	0	0-0	14	7-28	34	17-75	44	21-71	115.5	85-145
By referral type														
'Routine'	2030	24.4	13	1-42	1	0-7	20	8-42	45.5	20-82	75	42-141	113.5	62.5-207
'Other'	1122	19.1	5	0-26	0	0-5	14	7-29	35	13-77	63	34-83	89	54-180
'Emergency'	1642	25.9	4	0-21	0	0-2	8	3-19	21	11-46	35	16-71	56	30-128
'Two-week' referral	824	14.5	3	0-20	0	0-2	15	7-31	35	19-63	45	28-85	103.5	52.5-169.5

**Table 2.** Spearman's rank correlation co-efficient and ROC AUC values for the association between primary care interval and category of number of pre-referral consultations (n = 13 035)

	N	Spearman's r	Spearman's r P-value	ROC area under the curve <sup>a</sup>	ROC area 95% lower CI	ROC area 95% upper CI
All patients	13035	0.70	<0.0001	0.88	0.87	0.89
By cancer						
Breast	2367	0.56	<0.0001 <sup>b</sup>	0.96	0.93	0.98
Prostate	2201	0.55		0.84	0.82	0.86
Colorectal	1999	0.71		0.86	0.84	0.88
Lung	1421	0.62		0.83	0.81	0.86
Melanoma	735	0.61		0.87	0.81	0.94
Bladder	721	0.72		0.90	0.87	0.93
Lymphoma	590	0.70		0.87	0.84	0.90
Oesophageal	500	0.75		0.87	0.84	0.91
Endometrial	358	0.64		0.89	0.84	0.94
Leukaemia	358	0.59		0.81	0.76	0.87
Ovarian	324	0.68		0.86	0.82	0.91
Pancreatic	303	0.71		0.89	0.85	0.93
Renal	272	0.60		0.84	0.79	0.90
Stomach	237	0.76		0.88	0.83	0.92
Oropharyngeal	178	0.72		0.85	0.79	0.91
Multiple myeloma	176	0.73		0.88	0.83	0.93
Brain	159	0.77	0.92	0.87	0.96	
Unknown primary	136	0.76	0.87	0.81	0.93	
By referral type						
'Two-week' referral	8241	0.72	<0.0001 <sup>b</sup>	0.89	0.88	0.90
'Emergency'	1642	0.72		0.88	0.87	0.90
'Routine'	2030	0.67		0.85	0.83	0.87
'Other'	1122	0.67		0.88	0.85	0.90

Abbreviations: AUC = area under curve; CI = confidence interval; ROC = receiver-operator characteristic.

<sup>a</sup>For this analysis, the number of pre-referral consultations was dichotomised into 3+ vs 1–2 pre-referral consultation categories. Sensitivity analysis indicated that ROC AUC values were similar when other binary outcomes (e.g., '2+' vs 1, or '4+' vs 1, 2 or 3) were used, see Online Supplementary Information.<sup>b</sup>For any cancer type, and for any referral type.

A correlation between pre-referral consultations and duration of primary care interval might be intuitive but evidence about this association is sparse (Bjerager *et al*, 2006). The findings corroborate evidence suggesting that suspecting the diagnosis of cancer is generally more challenging for cancers without specific symptoms (e.g., multiple myeloma, stomach cancer) compared with cancers with specific signs and symptoms (breast cancer, melanoma) (Rubin *et al*, 2011; Lyratzopoulos *et al*, 2012).

Strengths of our study include its large sample size, the inclusion of patients with 18 different cancers and the use of a range of analytical approaches. There are several weaknesses. Some consultations may have been misclassified, either because of poor recording of presenting symptoms or misattribution of recorded symptoms—particularly in the context of co-morbidity (Weller *et al*, 2012). However, we judge that such errors are unlikely to have introduced substantial bias. Although there was no independent validation of the audit data, patients responding to the (2010) Cancer Patient Experience Survey reported similar patterns of variation by cancer in relation to the number of pre-referral consultations (Lyratzopoulos *et al*, 2012). Considering generalisability, although patients included in the audit reflect national incidence statistics in respect of cancer type, age and sex (Rubin

*et al*, 2011), it is conceivable that organisational factors and care quality in participating practices were different to other (non-participating) practices. Finally, we could not examine whether any of the prolonged values of either the primary care interval or the number of pre-referral consultations were either justifiable or preventable.

A more liberal policy for referral and investigation of patients with non-specific symptoms may increase the number of cancer patients diagnosed after one or two consultations; but at the expense of additional patient anxiety and healthcare utilisation costs for patients who will be investigated but found not to have cancer. Wider access by general practitioners to specialist diagnostic tests is increasingly being advocated (Department of Health, 2012). There is a need to monitor the impact of general practitioner-led investigations on the promptness of diagnosis of cancer (and other pathologies) and on resource use. This can be achieved by a programme of primary care audit that encompasses use of diagnostic imaging or endoscopic investigations. Point-Of-Care diagnostic technologies can also have a part in reducing the number of consultations before referral, and such tests merit further development and evaluation.

The findings support efforts to improve the timeliness of diagnosis by improving the sensitivity of the appraisal of cancer symptoms by general practitioners, for example, using clinical decision support tools (Hamilton, 2009; Hippisley-Cox and Coupland, 2012). Raising awareness of the importance of persistent symptoms among patients may also help reduce between-consultation intervals (and therefore improve timeliness of diagnosis), although such improvements may not necessarily reduce the number of pre-referral consultations. Research and policy initiatives can be further prioritised, focusing on patients with cancers that are more 'difficult-to-suspect' because of poor symptom specificity (e.g., multiple myeloma, lung, stomach and pancreatic cancer), which are typically associated with longer primary care intervals and greater number of pre-referral consultations (Lyratzopoulos *et al.*, 2012).

In conclusion, both the number of pre-referral consultations and primary care interval are important measures of the timeliness of cancer diagnosis and are inter-related. Improving the sensitivity of symptom appraisal by general practitioners to detect cancer symptoms should be prioritised by research and policy initiatives. Development and evaluation of interventions can particularly focus on patients with difficult-to-suspect cancers.

## ACKNOWLEDGEMENTS

We are grateful to all primary care professionals in participating practices for collecting, collating and submitting anonymous data; and respective Cancer Networks, the Royal College of General Practitioners and the National Cancer Action Team for supporting the audit. An integral part of the National Audit of Cancer Diagnosis in Primary Care project was that anonymous data may subsequently be used for purposes of early diagnosis research, and access to such data was granted by the Audit steering group. This paper is independent research arising from a Post-Doctoral Fellowship award to GL supported by the National Institute for Health Research. The views expressed in this publication are those of the authors and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health.

## REFERENCES

- Allgar VL, Neal RD (2005) Delays in the diagnosis of six cancers: analysis of data from the National Survey of NHS Patients: Cancer. *Br J Cancer* **92**(11): 1959–1970.
- Anonymous (2012) Cancer detection and primary care...revisited. *Lancet Oncol* **13**(6): 559.
- Bjerager M, Palshof T, Dahl R, Vedsted P, Olesen F (2006) Delay in diagnosis of lung cancer in general practice. *Br J Gen Pract* **56**: 863–868.
- Cancer Research UK (2009) About the National Awareness and Early Diagnosis Initiative <http://www.cancerresearchuk.org/cancer-info/spotcancerearly/naedi/AboutNAEDI/>. last accessed December, 2012.
- Department of Health (2012) Direct access to diagnostic tests for cancer: best practice referral pathways for general practitioners [http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH\\_133510](http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_133510). last accessed December, 2012.
- Hamilton W (2009) The CAPER studies: five case-control studies aimed at identifying and quantifying the risk of cancer in symptomatic primary care patients. *Br J Cancer* **101**(Suppl 2): S80–S86.
- Hansen RP, Vedsted P, Sokolowski I, Søndergaard J, Olesen F (2011) Time intervals from first symptom to treatment of cancer: a cohort study of 2,212 newly diagnosed cancer patients. *BMC Health Serv Res* **11**: 284.
- Hippisley-Cox J, Coupland C (2012) Identifying patients with suspected colorectal cancer in primary care: derivation and validation of an algorithm. *Br J Gen Pract* **62**(594): e29–e37.
- Lyratzopoulos G, Neal RD, Barbiere JM, Rubin GP, Abel GA (2012) Variation in number of general practitioner consultations before hospital referral for cancer: findings from the 2010 National Cancer Patient Experience Survey in England. *Lancet Oncol* **13**(4): 353–365.
- Neal RD, Allgar VL (2005) Sociodemographic factors and delays in the diagnosis of six cancers: analysis of data from the 'National Survey of NHS Patients: Cancer'. *Br J Cancer* **92**(11): 1971–1975.
- NHS (2011) National Cancer Patient Experience Survey Programme 2010 National Survey Report. [http://www.dh.gov.uk/prod\\_consum\\_dh/groups/dh\\_digitalassets/documents/digitalasset/dh\\_122520.pdf](http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_122520.pdf), last accessed December, 2012.
- Pancreatic Cancer UK (2011) Study for survival 2011. [http://www.pancreaticcancer.org.uk/media/100292/report\\_final\\_for\\_web.pdf](http://www.pancreaticcancer.org.uk/media/100292/report_final_for_web.pdf), last accessed December, 2012.
- Rarer Cancer Foundation (2011) Primary cause? An audit of the experience in primary care of rarer cancer patients. <http://www.rarercancers.org.uk/images/stories/cdf/p8and9/primary%20cause%20-%20final.pdf>, last accessed December, 2012.
- Richards MA, Westcombe AM, Love SB, Littlejohns P, Ramirez AJ (1999) Influence of delay on survival in patients with breast cancer: a systematic review. *Lancet* **353**(9159): 1119–1126.
- Rubin G, Elliott K, McPhail S. National Audit of Cancer Diagnosis in Primary Care. Royal College of General Practitioners (2011) [http://www.rcgp.org.uk/pdf/National\\_Audit\\_of\\_Cancer\\_Diagnosis\\_in\\_Primary-Care.pdf](http://www.rcgp.org.uk/pdf/National_Audit_of_Cancer_Diagnosis_in_Primary-Care.pdf). last accessed December, 2012.
- Teenage Cancer Trust (2012) 'One in four teenage cancer patients forced to visit GP at least four times before referral'. <http://www.teenagecancertrust.org/who-we-are/media-centre/press-releases/one-in-four-teenage-cancer-patients-forced-to-visit/>, last accessed December, 2012.
- The Roy Castle Lung Cancer Foundation (2011) Explaining variations in lung cancer in England. <http://www.roycastle.org/Resources/Roy%20Castle/Documents/PDF/Variations%20in%20lung%20cancer%20in%20England.pdf>, last accessed December, 2012.
- Thomson CS, Forman D (2009) Cancer survival in England and the influence of early diagnosis: what can we learn from recent EUROcare results? *Br J Cancer* **101**(Suppl 2): S102–S109.
- Tørring ML, Frydenberg M, Hansen RP, Olesen F, Hamilton W, Vedsted P (2011) Time to diagnosis and mortality in colorectal cancer: a cohort study in primary care. *Br J Cancer* **104**(6): 934–940.
- Weller D, Vedsted P, Rubin G, Walter FM, Emery J, Scott S, Campbell C, Andersen RS, Hamilton W, Olesen F, Rose P, Nafees S, van Rijswijk E, Hiom S, Muth C, Beyer M, Neal RD (2012) The Aarhus statement: improving design and reporting of studies on early cancer diagnosis. *Br J Cancer* **106**(7): 1262–1267.



This work is licensed under the Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>

Supplementary Information accompanies this paper on British Journal of Cancer website (<http://www.nature.com/bjc>)